

Achieving Max-Min Fairness for MU-MISO with Partial CSIT: A Multicast Assisted Transmission

Hamdi Joudeh* and Bruno Clerckx*†

* Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

† School of Electrical Engineering, Korea University, Seoul, Korea

Email: {hamdi.joudeh10, b.clerckx}@imperial.ac.uk

Abstract—We address the max-min fairness design problem for a MU-MISO system with partial Channel State Information (CSI) at the Base Station (BS), consisting of an imperfect channel estimate and statistical knowledge of the estimation error, and perfect CSI at the receivers. The objective is to maximize the minimum Average Rate (AR) among users subject to a transmit power constraint. An unconventional transmission scheme is adopted where the Base Station (BS) transmits a common message in addition to the conventional private messages. In situations where the CSIT is not accurate enough to perform interference nulling, individual rates are assisted by allocating parts of the common message to different users according to their needs. The AR problem is transformed into an augmented Average Weighted Mean Square Error (AWMSE) problem, solved using Alternating Optimization (AO). The benefits of incorporating the common message are demonstrated through simulations.

Index Terms—MISO-JMB, max-min fairness, AWMSE.

I. INTRODUCTION

We consider a Multi-User (MU) Multiple Input Single Output (MISO) system where a Base Station (BS) equipped with N_t antennas serves $K \leq N_t$ single-antenna receivers. For a given transmit power P_t , the objective is to design precoders (or beamformers) that maximize the minimum rate among users; a criteria known in literature as max-min fairness [1]–[3]. In practical scenarios, perfect Channel State Information at the Transmitter (CSIT) is unguaranteed due to the varying nature of wireless channels and the limited resources available for CSIT acquisition. This prompted a number of robust fairness-based designs [4]–[6]. In this paper, we consider the case where the BS obtains an imperfect estimate of the channel (through feedback or training) and statistical properties of the estimation error, while users estimate and track their channels accurately. The problem is formulated in terms of the Average Rate (AR) which captures the overall performance w.r.t channel estimation errors. Contrary to our previous work in [6] which considers a similar scenario, a rather unorthodox transmission scheme is considered in this paper. In particular, the BS transmits a common message in addition to the conventionally transmitted private messages. This scheme is termed Joint Multicasting and Broadcasting (JMB).

JMB was employed in [7] to boost the achievable Degrees of Freedom (DoF) under estimation errors that decay with increased Signal to Noise Ratio (SNR). Assuming fixed noise variances at the receivers, the decaying power of the estimation

error scales as $O(P_t^{-\alpha})$, where $\alpha \geq 0$ is an exponent that represents the CSIT quality. This power relationship can be used to express a variety of imperfect CSIT scenarios [8]. For example, $\alpha = 0$ represents the case where the error power is fixed over the whole SNR regime, e.g. constant number of feedback bits. On the other hand, $\alpha = \infty$ corresponds to perfect CSIT. In DoF analysis, it is customary to truncate the exponent such that $\alpha \leq 1$, where $\alpha = 1$ is considered perfect from a DoF perspective as the error impact becomes negligible for $P_t \rightarrow \infty$. Under this assumption, conventional MU transmission that employs Zero Forcing (ZF) Beamforming (BF) achieves a sum DoF of $K\alpha$. Interestingly, the same sum DoF can be achieved using only part of the power given as P_t^α . This can be explained as follows: in the presence of CSIT uncertainty, perfect interference nulling cannot be achieved, and any DoF gain achieved by increasing the power above a certain threshold (i.e. P_t^α) is nullified by the increased interference. Nevertheless, the remaining part (i.e. $P_t - P_t^\alpha$) can be used to transmit a common message which achieves a DoF of $1-\alpha$ by decoding it while treating the private messages as noise. As the common message is decodable by all users, it can be cancelled from their received signals, and the DoF achieved by the private messages is unaffected.

Contribution and Organization: We employ JMB to maximize the minimum AR among users, where a given user's AR consists of a private AR, in addition to a portion of the common AR. This is achieved by jointly optimizing the precoders and the partition factors which determine how much of the common message is allocated to each user. This non-convex problem is transformed into an equivalent augmented Average Weighted Mean Square Error (AWMSE) problem, which is solved using Alternating Optimization (AO). Simulation results show the superiority of the proposed scheme over conventional MU transmission, which does not incorporate the common message. For example, while conventional transmission with $\alpha = 0$ yields flat rates in the high SNR regime [6], JMB employs the multicast part to achieve a non-vanishing DoF. To the best of our knowledge, multicast assisted robust fairness-based designs have not appeared in literature.

The rest of the paper is organized as follows: the system model and problem formulation are given in Section II. The equivalent AWMSE problem is formulated in Section III. The AO algorithm is proposed in Section IV. Simulation results are given in Section V, and Section VI concludes the paper.

Notation: Boldface uppercase letters denote matrices, boldface lowercase letters denote column vectors and standard letters denote scalars. $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate-transpose (Hermitian) operators, respectively. $\text{tr}(\cdot)$ and $\|\cdot\|$ are the trace and Euclidian norm operators, respectively. $E_x\{\cdot\}$ is the expectation w.r.t the random variable x .

II. SYSTEM MODEL AND PROBLEM FORMULATION

For the MISO setup described in the previous section, the vector of complex data symbols is given as $\mathbf{s} \triangleq [s_c, s_1, \dots, s_K]^T \in \mathbb{C}^{K+1}$, where s_c is the common symbol decodable by all users, s_i is a private symbol intended solely for the i th user, $i \in \mathcal{K}$ and $\mathcal{K} \triangleq \{1, \dots, K\}$. Entries of \mathbf{s} have zero-means, unity powers and are mutually uncorrelated such that $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$. \mathbf{s} is linearly precoded as

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{i=1}^K \mathbf{p}_i s_i \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{N_t}$ is the transmit vector. $\mathbf{p}_c \in \mathbb{C}^{N_t}$ and $\mathbf{p}_i \in \mathbb{C}^{N_t}$ are precoders for the common symbol and the i th private symbol respectively, from which $\mathbf{P} \triangleq [\mathbf{p}_c, \mathbf{p}_1, \dots, \mathbf{p}_K]$ is composed. The BS transmit power constraint is given as $E\{\mathbf{x}^H \mathbf{x}\} = \text{tr}(\mathbf{P}\mathbf{P}^H) \leq P_t$. The k th received signal is given as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k \quad (2)$$

where $\mathbf{h}_k \in \mathbb{C}^{N_t}$ is the channel impulse response vector between the BS and the k th user, from which the composite channel is defined as $\mathbf{H} \triangleq [\mathbf{h}_1, \dots, \mathbf{h}_K]$. $n_k \sim \mathcal{CN}(0, \sigma_{n_k}^2)$ is the AWGN. Throughout the paper, it is assumed that noise variances are equal across all users, i.e. $\sigma_{n_k}^2 = \sigma_n^2, \forall k \in \mathcal{K}$.

A. CSIT Uncertainty

Each of the K links exhibits independent fading, and remains almost constant over a frame of symbols, enabling users to estimate their channel vectors with high accuracy. On the other hand, CSIT experiences uncertainty arising from limited feedback, delays or mismatches. \mathbf{H} is written as a sum of the transmitter-side channel estimate $\widehat{\mathbf{H}} \triangleq [\widehat{\mathbf{h}}_1, \dots, \widehat{\mathbf{h}}_K]$ and the channel estimation error $\widetilde{\mathbf{H}} \triangleq [\widetilde{\mathbf{h}}_1, \dots, \widetilde{\mathbf{h}}_K]$, such that $\mathbf{H} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}$. The CSIT consists of $\widehat{\mathbf{H}}$, in addition to some statistical knowledge of $\widetilde{\mathbf{H}}$. Particularly, the BS knows the probability distribution of the actual channel given the available estimate, i.e. $f_{\mathbf{H}|\widehat{\mathbf{H}}}(\mathbf{H}) = f_{\widetilde{\mathbf{H}}}(\mathbf{H} - \widehat{\mathbf{H}})$.

B. MSE and MMSE

The k th user obtains an estimate of the common symbol by applying a scalar equalizer $g_{c,k}(\mathbf{h}_k)$ to (2) such that $\widehat{s}_{c,k} = g_{c,k}(\mathbf{h}_k)y_k$. Assuming that the common symbol is successfully decoded by all users, the common symbol's receive signal part is reconstructed and cancelled from y_k . This improves the detectability of s_k , which is then estimated by applying $g_k(\mathbf{h}_k)$ such that $\widehat{s}_k = g_k(\mathbf{h}_k)(y_k - \mathbf{h}_k^H \mathbf{p}_c s_{c,k})$. The notations $g_{c,k}(\mathbf{h}_k)$ and $g_k(\mathbf{h}_k)$ are used to emphasise the dependencies on the actual channel, as each user is assumed to have perfect knowledge of its own channel vector. (\mathbf{h}_k) is omitted for brevity unless special emphasis is necessary. This is also used

with other variables that depend on the actual channel. For the k th user, the MSEs defined as $\varepsilon_{c,k} \triangleq E_{\mathbf{s},n_k}\{\|\widehat{s}_{c,k} - s_c\|^2\}$ and $\varepsilon_k \triangleq E_{\mathbf{s},n_k}\{\|\widehat{s}_k - s_k\|^2\}$ are given as

$$\varepsilon_{c,k}(\mathbf{h}_k) = |g_{c,k}|^2 T_{c,k} - 2\Re\{g_{c,k} \mathbf{h}_k^H \mathbf{p}_c\} + 1 \quad (3a)$$

$$\varepsilon_k(\mathbf{h}_k) = |g_k|^2 T_k - 2\Re\{g_k \mathbf{h}_k^H \mathbf{p}_k\} + 1 \quad (3b)$$

where $T_{c,k} = |\mathbf{p}_c^H \mathbf{h}_k|^2 + T_k$ and $T_k = \sum_{i=1}^K |\mathbf{p}_i^H \mathbf{h}_k|^2 + \sigma_n^2$. Optimum $g_{c,k}$ and g_k are obtained from $\frac{\partial \varepsilon_{c,k}}{\partial g_{c,k}} = 0$ and $\frac{\partial \varepsilon_k}{\partial g_k} = 0$, yielding the well-known Minimum MSE (MMSE) equalizers:

$$g_{c,k}^{\text{MMSE}}(\mathbf{h}_k) = \mathbf{p}_c^H \mathbf{h}_k T_{c,k}^{-1} \text{ and } g_k^{\text{MMSE}}(\mathbf{h}_k) = \mathbf{p}_k^H \mathbf{h}_k T_k^{-1}. \quad (4)$$

Substituting (4) into (3), the k th user's MMSEs are given as

$$\varepsilon_{c,k}^{\text{MMSE}}(\mathbf{h}_k) = T_{c,k}^{-1} E_{c,k} \text{ and } \varepsilon_k^{\text{MMSE}}(\mathbf{h}_k) = T_k^{-1} E_k \quad (5)$$

where $E_{c,k} = T_{c,k} - |\mathbf{p}_c^H \mathbf{h}_k|^2 = T_k$ and $E_k = T_k - |\mathbf{p}_k^H \mathbf{h}_k|^2$.

C. Achievable Rate and Average Rate

The MMSE and the Signal to Interference plus Noise Ratio (SINR) are related such that $\gamma_{c,k} = (1 - \varepsilon_{c,k}^{\text{MMSE}})/\varepsilon_{c,k}^{\text{MMSE}}$ and $\gamma_k = (1 - \varepsilon_k^{\text{MMSE}})/\varepsilon_k^{\text{MMSE}}$, where $\gamma_{c,k}$ and γ_k are the k th user's SINRs. Therefore, the k th user's maximum achievable common rate and private rate are written as $R_{c,k}(\mathbf{h}_k) = -\log_2(\varepsilon_{c,k}^{\text{MMSE}})$ and $R_k(\mathbf{h}_k) = -\log_2(\varepsilon_k^{\text{MMSE}})$, respectively. The common message is transmitted at a common rate defined as $R_c \triangleq \min_j \{R_{c,j}\}_{j=1}^K$, which ensures that it is decodable by all users. As the multicast part is used to boost the individual rates achieved by users, the k th user is allocated a fraction c_k of the common message, where $c_k \geq 0$ and $\sum_{k=1}^K c_k = 1$. The corresponding portion of the common rate is given as $c_k R_c$. In this case, the k th user's total achievable rate is given as $R_k + c_k R_c$. In a block-based transmission, this can be achieved by dedicating different parts of the common data block to different users, where any fraction c_k can be achieved for sufficiently long blocks. To achieve fairness among users, the objective would be to design \mathbf{P} and $\mathbf{c} \triangleq [c_1, \dots, c_K]^T$ that maximize the minimum total rate among users defined as $R \triangleq \min_k \{R_k + c_k R_c\}_{k=1}^K$. However, rates depend on the actual channel, and hence cannot be considered as optimization metrics at the BS. Alternatively, we consider the Average Rates (ARs) defined as: $E_{\mathbf{h}_k|\widehat{\mathbf{h}}_k}\{R_{c,k}\}$ and $E_{\mathbf{h}_k|\widehat{\mathbf{h}}_k}\{R_k\}$, where averaging is taken over the CSIT error. In the following, $E_{\mathbf{h}_k|\widehat{\mathbf{h}}_k}\{\cdot\}$ will be simply referred to as $E\{\cdot\}$.

D. Sample Average Function

In order to formulate a deterministic problem, the stochastic ARs are approximate by corresponding Sample Average Functions (SAFs) obtained by taking the ensemble average over a sample of M independent identically distributed (i.i.d) realizations drawn from the distribution $f_{\mathbf{H}|\widehat{\mathbf{H}}}$. The sample is defined as $\mathbf{H}_{\mathcal{M}} \triangleq \{\mathbf{H}^{(m)} \mid m \in \mathcal{M}\}$, where $\mathbf{H}^{(m)} \triangleq [\mathbf{h}_1^{(m)}, \dots, \mathbf{h}_K^{(m)}]$ is the m th Monte-Carlo realization, and $\mathcal{M} \triangleq \{1, \dots, M\}$. The SAFs are given as: $\bar{R}_{c,k}^{(M)} = \frac{1}{M} \sum_{m=1}^M R_{c,k}^{(m)}$ and $\bar{R}_k^{(M)} = \frac{1}{M} \sum_{m=1}^M R_k^{(m)}$, where $R_{c,k}^{(m)} \triangleq R_{c,k}(\mathbf{h}_k^{(m)})$ and $R_k^{(m)} \triangleq R_k(\mathbf{h}_k^{(m)})$ are the rates associated with the realization $\mathbf{h}_k^{(m)}$. In the following, the

superscript (m) is used to indicate the association of variables with the m th Monte-Carlo realization. It should be noted that \mathbf{P} is fixed over the M realizations of the rates, which follows from the definition of the ARs. This also reflects the fact that \mathbf{P} is optimized at the BS using partial CSI knowledge.

Assumption 1. In the following, we assume that $\frac{\sigma_n^2}{\|\mathbf{h}_k\|^2 P_t} > 0$ with probability 1, $\forall k \in \mathcal{K}$.

Alternatively, we can say that $\text{SNR} = P_t/\sigma_n^2$ can only grow finitely large, and channel gains are finite. Assumption 1 yields $\varepsilon_{c,k}^{\text{MMSE}}, \varepsilon_k^{\text{MMSE}} > 0$ with probability 1, as the presence of a nonzero noise variance dictates that $E_{c,k}, E_k > 0$. This also implies that rates are finite, and by the strong law of large numbers we can write

$$\bar{R}_{c,k} \triangleq \lim_{M \rightarrow \infty} \bar{R}_{c,k}^{(M)} = E\{R_{c,k}\}, \text{almost surely} \quad (6a)$$

$$\bar{R}_k \triangleq \lim_{M \rightarrow \infty} \bar{R}_k^{(M)} = E\{R_k\}, \text{almost surely} \quad (6b)$$

where $\bar{R}_{c,k}$ and \bar{R}_k are the approximated ARs for a sufficiently large M , which will be simply referred to as the ARs. Moreover, the common AR is defined as $\bar{R}_c \triangleq \min_j \{\bar{R}_{c,j}\}_{j=1}^K$.

E. Problem Formulation

The minimum total AR among the K users is defined as $\bar{R} \triangleq \min_k \{\bar{R}_k + c_k \bar{R}_c\}_{k=1}^K$. The objective is to design \mathbf{P} and \mathbf{c} such that \bar{R} is maximized. This problem is formulated as

$$\mathcal{R} : \max_{\bar{R}, \bar{R}_c, \mathbf{P}, \mathbf{c}} \bar{R} \quad (7a)$$

$$\text{s.t. } \bar{R}_k + c_k \bar{R}_c \geq \bar{R}, \forall k \in \mathcal{K} \quad (7b)$$

$$\bar{R}_{c,k} \geq \bar{R}_c, \forall k \in \mathcal{K} \quad (7c)$$

$$c_k \geq 0, \forall k \in \mathcal{K} \quad (7d)$$

$$\sum_{k=1}^K c_k = 1 \quad (7e)$$

$$\text{tr}(\mathbf{P} \mathbf{P}^H) \leq P_t \quad (7f)$$

where the constraints in (7b) and (7c) are introduced to eliminate the potential non-smoothness arising from the pointwise minimizations in \bar{R}_c and \bar{R} . \mathcal{R} is non-convex and appears to be very challenging to solve.

III. EQUIVALENT AWMSE OPTIMIZATION PROBLEM

In this section, \mathcal{R} is transformed into a more tractable equivalent problem by exploiting the relationship between rates and augmented WMSEs, an approach inspired from [3], [9]. We start by introducing the main components used to construct the equivalent problem, i.e. the augmented WMSEs:

$$\xi_{c,k}(\mathbf{h}_k, g_{c,k}, u_{c,k}) = u_{c,k}(\mathbf{h}_k) \varepsilon_{c,k}(\mathbf{h}_k) - \log_2(u_{c,k}(\mathbf{h}_k)) \quad (8a)$$

$$\xi_k(\mathbf{h}_k, g_k, u_k) = u_k(\mathbf{h}_k) \varepsilon_k(\mathbf{h}_k) - \log_2(u_k(\mathbf{h}_k)) \quad (8b)$$

where $u_{c,k}(\mathbf{h}_k) \geq 0$ and $u_k(\mathbf{h}_k) \geq 0$ are weights associated with the k th user's MSEs. The dependencies of $\xi_{c,k}$ and ξ_k on different variables are highlighted in (8) for their significance in the following analysis, where we establish the following WMSE-Rate relationship:

$$\min_{u_{c,k}, g_{c,k}} \xi_{c,k} = 1 - R_{c,k} \quad \text{and} \quad \min_{u_k, g_k} \xi_k = 1 - R_k. \quad (9)$$

This can be shown as follows. From $\frac{\partial \xi_{c,k}}{\partial g_{c,k}} = 0$ and $\frac{\partial \xi_k}{\partial g_k} = 0$, the optimum equalizers are obtained as $g_{c,k}^* = g_{c,k}^{\text{MMSE}}$ and $g_k^* = g_k^{\text{MMSE}}$. Substituting this back into (8), we obtain the augmented WMSEs written as

$$\xi_{c,k}^{\text{MMSE}}(\mathbf{h}_k, u_{c,k}) = u_{c,k} \varepsilon_{c,k}^{\text{MMSE}} - \log_2(u_{c,k}) \quad (10a)$$

$$\xi_k^{\text{MMSE}}(\mathbf{h}_k, u_k) = u_k \varepsilon_k^{\text{MMSE}} - \log_2(u_k). \quad (10b)$$

Furthermore, from $\frac{\partial \xi_{c,k}^{\text{MMSE}}}{\partial u_{c,k}} = 0$ and $\frac{\partial \xi_k^{\text{MMSE}}}{\partial u_k} = 0$, we obtain the optimum MMSE weights: $u_{c,k}^* = u_{c,k}^{\text{MMSE}} \triangleq (\varepsilon_{c,k}^{\text{MMSE}})^{-1}$ and $u_k^* = u_k^{\text{MMSE}} \triangleq (\varepsilon_k^{\text{MMSE}})^{-1}$. Substituting this back into (10) yields the relationship in (9). It is evident from (5) that the MMSE weights are dependent on the channel.

The equivalent problem is formulated using the augmented AWMSEs defined as: $E\{\xi_{c,k}\}$ and $E\{\xi_k\}$. Before we proceed, the augmented AWMSEs are approximated as:

$$\bar{\xi}_{c,k}^{(M)} = \frac{1}{M} \sum_{m=1}^M \xi_{c,k}^{(m)} \quad \text{and} \quad \bar{\xi}_k^{(M)} = \frac{1}{M} \sum_{m=1}^M \xi_k^{(m)}, \text{ where}$$

$$\xi_{c,k}^{(m)} \triangleq \xi_{c,k}(\mathbf{h}_k^{(m)}, g_{c,k}^{(m)}, u_{c,k}^{(m)}) \quad \text{and} \quad \xi_k^{(m)} \triangleq \xi_k(\mathbf{h}_k^{(m)}, g_k^{(m)}, u_k^{(m)})$$

correspond to the m th realization of the augmented WMSEs, which depend on the m th realization of the equalizers: $g_{c,k}^{(m)} \triangleq g_{c,k}(\mathbf{h}_k^{(m)})$ and $g_k^{(m)} \triangleq g_k(\mathbf{h}_k^{(m)})$, and the weights: $u_{c,k}^{(m)} \triangleq u_{c,k}(\mathbf{h}_k^{(m)})$ and $u_k^{(m)} \triangleq u_k(\mathbf{h}_k^{(m)})$. For compactness, we define the set of equalizers associated with the M realizations and the K users as: $\mathbf{G} \triangleq \{\mathbf{g}_{c,k}, \mathbf{g}_k \mid k \in \mathcal{K}\}$, where $\mathbf{g}_{c,k} \triangleq \{g_{c,k}^{(m)} \mid m \in \mathcal{M}\}$ and $\mathbf{g}_k \triangleq \{g_k^{(m)} \mid m \in \mathcal{M}\}$. In a similar manner, we define: $\mathbf{U} \triangleq \{\mathbf{u}_{c,k}, \mathbf{u}_k \mid k \in \mathcal{K}\}$, where $\mathbf{u}_{c,k} \triangleq \{u_{c,k}^{(m)} \mid m \in \mathcal{M}\}$ and $\mathbf{u}_k \triangleq \{u_k^{(m)} \mid m \in \mathcal{M}\}$. The approximated augmented AWMSEs for a sufficiently large M are defined as $\bar{\xi}_{c,k} \triangleq \lim_{M \rightarrow \infty} \bar{\xi}_{c,k}^{(M)}$ and $\bar{\xi}_k \triangleq \lim_{M \rightarrow \infty} \bar{\xi}_k^{(M)}$, which will be simply referred to as the AWMSEs. The same approach used to prove (9) can be employed to show that

$$\min_{\mathbf{u}_{c,k}, \mathbf{g}_{c,k}} \bar{\xi}_{c,k} = 1 - \bar{R}_{c,k} \quad \text{and} \quad \min_{\mathbf{u}_k, \mathbf{g}_k} \bar{\xi}_k = 1 - \bar{R}_k. \quad (11)$$

where optimality conditions are checked separately for each realization. The sets of optimum MMSE equalizers associated with (11) are defined as $\mathbf{g}_{c,k}^{\text{MMSE}} \triangleq \{g_{c,k}^{\text{MMSE}(m)} \mid m \in \mathcal{M}\}$ and $\mathbf{g}_k^{\text{MMSE}} \triangleq \{g_k^{\text{MMSE}(m)} \mid m \in \mathcal{M}\}$. In the same manner, the sets of optimum MMSE weights are defined as $\mathbf{u}_{c,k}^{\text{MMSE}(m)} \triangleq \{u_{c,k}^{\text{MMSE}(m)} \mid m \in \mathcal{M}\}$ and $\mathbf{u}_k^{\text{MMSE}(m)} \triangleq \{u_k^{\text{MMSE}(m)} \mid m \in \mathcal{M}\}$. For the K users, the MMSE solution is composed as $\mathbf{G}^{\text{MMSE}} \triangleq \{\mathbf{g}_{c,k}^{\text{MMSE}}, \mathbf{g}_k^{\text{MMSE}} \mid k \in \mathcal{K}\}$ and $\mathbf{U}^{\text{MMSE}} \triangleq \{\mathbf{u}_{c,k}^{\text{MMSE}}, \mathbf{u}_k^{\text{MMSE}} \mid k \in \mathcal{K}\}$.

A. Augmented AWMSE Minimization

Prompted by the relationship in (11), the augmented AWMSE minimization problem is formulated as

$$\begin{aligned}
\mathcal{A} : & \min_{\bar{\xi}, \bar{\xi}_c, \mathbf{P}, \mathbf{c}, \mathbf{U}, \mathbf{G}} \bar{\xi} & (12a) \\
\text{s.t. } & \bar{\xi}_k + c_k(\bar{\xi}_c - 1) \leq \bar{\xi}, \forall k \in \mathcal{K} & (12b) \\
& \bar{\xi}_{c,k} \leq \bar{\xi}_c, \forall k \in \mathcal{K} & (12c) \\
& c_k \geq 0, \forall k \in \mathcal{K} & (12d) \\
& \sum_{k=1}^K c_k = 1 & (12e) \\
& \text{tr}(\mathbf{P}\mathbf{P}^H) \leq P_t & (12f)
\end{aligned}$$

where $(\bar{\xi}, \bar{\xi}_c)$ are auxiliary variables. The equivalence between \mathcal{A} under the MMSE solution, and \mathcal{R} is demonstrated as follows. Substituting $(\mathbf{g}_k^{\text{MMSE}}, \mathbf{u}_k^{\text{MMSE}})$ and $(\mathbf{g}_{c,k}^{\text{MMSE}}, \mathbf{u}_{c,k}^{\text{MMSE}})$ into (12b) and (12c) respectively, and rearranging, we obtain

$$\bar{R}_k + c_k(1 - \bar{\xi}_c) \geq 1 - \bar{\xi}, \forall k \in \mathcal{K} \quad (13a)$$

$$\bar{R}_{c,k} \geq 1 - \bar{\xi}_c, \forall k \in \mathcal{K}. \quad (13b)$$

It can be seen that (13a) and (13b) are equivalent to (7b) and (7c) respectively, where

$$\bar{R} = 1 - \bar{\xi} \quad \text{and} \quad \bar{R}_c = 1 - \bar{\xi}_c. \quad (14)$$

Moreover, the monotonicity of the relationships in (14) implies that the two problems are equivalent under the MMSE solution, i.e. maximizing \bar{R} is equivalent to minimizing $\bar{\xi}$. This relationship is exploited in the following section.

IV. ALTERNATING OPTIMIZATION ALGORITHM

Although \mathcal{A} is non-convex in the joint set of optimization variables, it is convex in each of the blocks \mathbf{G} , \mathbf{U} , \mathbf{c} and \mathbf{P} . We propose an AO algorithm that exploits this block-wise convexity. Each iteration of the algorithm consists of three steps: 1) updating \mathbf{G} and \mathbf{U} , 2) updating \mathbf{c} , 3) updating \mathbf{P} .

A. Updating the Equalizers and Weights

In n th iteration of the AO algorithm, the equalizers and weights are updated as: $\mathbf{G} = \mathbf{G}^{\text{MMSE}}(\ddot{\mathbf{P}})$ and $\mathbf{U} = \mathbf{U}^{\text{MMSE}}(\ddot{\mathbf{P}})$, where $\ddot{\mathbf{P}}$ is the precoding matrix obtained from $(n-1)$ th iteration. For the resulting point $(\ddot{\mathbf{P}}, \mathbf{G}^{\text{MMSE}}(\ddot{\mathbf{P}}), \mathbf{U}^{\text{MMSE}}(\ddot{\mathbf{P}}))$, the updated AWMSEs in (12) are denoted by $\ddot{\xi}_{c,k}$ and $\ddot{\xi}_k$. Moreover, the common AWMSE is given as $\ddot{\xi}_c = \min_j \{\ddot{\xi}_{c,j}\}_{j=1}^K$. Finally, the cost function in (12) is reduced by minimizing the individual AWMSEs.

B. Updating Partition Coefficients

Using the updated AWMSEs from the previous subsection, and employing the relationships in (13) and (14), the problem of optimizing \mathbf{c} is formulated in terms of the ARs as

$$\mathcal{A}_c : \min_{\bar{R}, \mathbf{c}} -\bar{R} \quad (15a)$$

$$\text{s.t. } \bar{R}_k + c_k \ddot{R}_c \geq \bar{R}, \forall k \in \mathcal{K} \quad (15b)$$

$$\sum_{k=1}^K c_k = 1 \quad (15c)$$

$$c_k \geq 0, \forall k \in \mathcal{K} \quad (15d)$$

where $\ddot{R}_k = 1 - \ddot{\xi}_k$ and $\ddot{R}_c = 1 - \ddot{\xi}_c$. While optimizing \mathbf{c} does not influence the private ARs or the common AR,

it redistributes the common AR among users in a way that further reduces the cost function of \mathcal{A} . A solution for problem \mathcal{A}_c is developed by examining the KKT optimality conditions. The Lagrangian of (15) is given as

$$L = -\bar{R} + \sum_{k=1}^K \mu_k (\bar{R} - \ddot{R}_k - c_k \ddot{R}_c) + \lambda \left(\sum_{k=1}^K c_k - 1 \right) - \sum_{k=1}^K \nu_k c_k$$

where $\boldsymbol{\mu} \triangleq \{\mu_k \mid k \in \mathcal{K}\}$, λ , and $\boldsymbol{\nu} \triangleq \{\nu_k \mid k \in \mathcal{K}\}$ are the multipliers associated with the constraints in (15a), (15b), and (15c), respectively. The KKT conditions can be written as

$$\frac{\partial L}{\partial \bar{R}} = 0 \Rightarrow \sum_{k=1}^K \mu_k = 1 \quad (16a)$$

$$\frac{\partial L}{\partial c_k} = 0 \Rightarrow -\mu_k \ddot{R}_c + \lambda - \nu_k = 0, \forall k \in \mathcal{K} \quad (16b)$$

$$\mu_k \geq 0, \mu_k (\bar{R} - \ddot{R}_k - c_k \ddot{R}_c) = 0, \forall k \in \mathcal{K} \quad (16c)$$

$$\nu_k \geq 0, \nu_k c_k = 0, \forall k \in \mathcal{K}. \quad (16d)$$

For $\ddot{R}_c = 0$, optimizing \mathbf{c} is irrelevant. On the other hand, $\ddot{R}_c > 0$ implies that $\lambda > 0$. This is concluded from (16b) which can be rewritten as $\lambda = \nu_k + \mu_k \ddot{R}_c$, (16a) which dictates that at least one element in $\boldsymbol{\mu}$ is nonzero, and the positivity of ν_k from (16d). Now we consider c_k under three cases:

1) For the first case, we assume $\bar{R} < \ddot{R}_k$. In this case, (16c) dictates that $\mu_k = 0$, as $\bar{R} - \ddot{R}_k - c_k \ddot{R}_c = 0$ is impossible due to the positivity of $c_k \ddot{R}_c$. Moreover, we have $\lambda = \nu_k > 0$ from (16b). This requires $c_k = 0$ as seen from (16d). In other words, if a given user's private AR is higher than the optimum \bar{R} , it will not be allocated a fraction of the common AR. Moreover, the corresponding constraints in (15b) are inactive.

2) Next, we consider $\bar{R} = \ddot{R}_k$. Assuming that $c_k > 0$, then we must have $\mu_k = 0$ to satisfy (16c). This implies that $\lambda = \nu_k > 0$, which contradicts (16d). Therefore, we must have $c_k = 0$ and $\bar{R} = \ddot{R}_k$, i.e. the corresponding constraints in (15b) are active.

3) Finally, we consider $\bar{R} > \ddot{R}_k$. Assuming that $\mu_k = 0$, then we have $\lambda = \nu_k > 0$ and $c_k = 0$. However, this is an infeasible solution as it contradicts (15b). Therefore, we must have $\mu_k > 0$ and $\bar{R} = \ddot{R}_k + c_k \ddot{R}_c$. For this case, the corresponding constraints in (15b) are also active.

From the previous analysis, we can write

$$c_k = \max \{0, \ddot{R}_c^{-1}(\bar{R} - \ddot{R}_k)\}, \forall k \in \mathcal{K} \quad (17)$$

which suggests that the optimum \mathbf{c} can be calculated using a form of water-filling, where \bar{R} is the water level. Without loss of generality, we assume that the private ARs are ordered in an ascending manner, i.e. $\ddot{R}_i \geq \ddot{R}_j, \forall i, j \in \mathcal{K}$ and $i > j$. Assuming that all constraints in (15b) are active, and using the constraint in (15c), the water level is calculated as $\bar{R} = K^{-1}(\ddot{R}_c + \sum_{k=1}^K \ddot{R}_k)$, from which we obtain $c_k = \ddot{R}_c^{-1}(\bar{R} - \ddot{R}_k), \forall k \in \mathcal{K}$. If the K th user (with the greatest private AR) has $c_K < 0$, then the first case of the three applies. The K th user is discarded and \bar{R} and \mathbf{c} are recalculated for the remaining $K-1$ users. Otherwise, the optimum solution is obtained. This is summarized in Algorithm 1.

Algorithm 1 Optimizing the partition coefficients.

- 1: **Initialize:** $K_c \leftarrow K + 1$
- 2: **repeat**
- 3: $K_c \leftarrow K_c - 1$ and $\mathcal{K}_c \leftarrow \{1, \dots, K_c\}$
- 4: $\bar{R} \leftarrow K_c^{-1}(\ddot{R}_c + \sum_{k=1}^{K_c} \ddot{R}_k)$
- 5: $c_k \leftarrow \ddot{R}_c^{-1}(\bar{R} - \ddot{R}_k), \forall k \in \mathcal{K}_c$
- 6: $c_k \leftarrow 0, \forall k \in \mathcal{K} \setminus \mathcal{K}_c$
- 7: **until** $c_{K_c} \geq 0$

C. Updating the Precoders

To formulate the problem of updating the precoders, the AWMSEs are written in terms of the updated blocks \mathbf{G} and \mathbf{U} , and the block \mathbf{P} which is yet to be updated. For this purpose, we introduce the AWMMSE-components: $\bar{\Psi}_{c,k}$, $\bar{\Psi}_k$, $\bar{t}_{c,k}$, \bar{t}_k , $\bar{\mathbf{f}}_k$, $\bar{\mathbf{f}}_{c,k}$, \bar{u}_k , $\bar{u}_{c,k}$ and \bar{v}_k , which are obtained using the updated \mathbf{G} and \mathbf{U} . In particular, $\bar{u}_{c,k}$ and \bar{u}_k are calculated by taking the ensemble averages over the M realizations of $u_{c,k}^{(m)}$ and $u_k^{(m)}$. The rest are calculated in a similar manner by averaging over their corresponding realizations:

$$\begin{aligned} t_{c,k}^{(m)} &= u_{c,k}^{(m)} |g_{c,k}^{(m)}|^2 & \text{and} & \quad t_k^{(m)} = u_k^{(m)} |g_k^{(m)}|^2 \\ \Psi_{c,k}^{(m)} &= t_{c,k}^{(m)} \mathbf{h}_k^{(m)} \mathbf{h}_k^{(m)H} & \text{and} & \quad \Psi_k^{(m)} = t_k^{(m)} \mathbf{h}_k^{(m)} \mathbf{h}_k^{(m)H} \\ \mathbf{f}_{c,k}^{(m)} &= u_{c,k}^{(m)} \mathbf{h}_k^{(m)} g_{c,k}^{(m)H} & \text{and} & \quad \mathbf{f}_k^{(m)} = u_k^{(m)} \mathbf{h}_k^{(m)} g_k^{(m)H} \\ v_{c,k}^{(m)} &= \log_2(u_{c,k}^{(m)}) & \text{and} & \quad v_k^{(m)} = \log_2(u_k^{(m)}) . \end{aligned}$$

Using the AWMMSE-components, the AWMSEs are given as

$$\begin{aligned} \bar{\xi}_{c,k} &= \mathbf{p}_c^H \bar{\Psi}_{c,k} \mathbf{p}_c + \sum_{i=1}^K \mathbf{p}_i^H \bar{\Psi}_{c,k} \mathbf{p}_i + \sigma_n^2 \bar{t}_{c,k} - 2\Re\{\bar{\mathbf{f}}_{c,k}^H \mathbf{p}_c\} \\ &\quad + \bar{u}_{c,k} - \bar{v}_{c,k} \end{aligned} \quad (18a)$$

$$\bar{\xi}_k = \sum_{i=1}^K \mathbf{p}_i^H \bar{\Psi}_k \mathbf{p}_i + \sigma_n^2 \bar{t}_k - 2\Re\{\bar{\mathbf{f}}_k^H \mathbf{p}_k\} + \bar{u}_k - \bar{v}_k \quad (18b)$$

from which the problem of optimizing \mathbf{P} is formulated as

$$\mathcal{A}_{\mathbf{P}} : \min_{\bar{\xi}, \bar{\xi}_c, \mathbf{P}} \bar{\xi} \quad (19a)$$

$$\begin{aligned} \text{s.t. } & \sum_{i=1}^K \mathbf{p}_i^H \bar{\Psi}_k \mathbf{p}_i + \sigma_n^2 \bar{t}_k - 2\Re\{\bar{\mathbf{f}}_k^H \mathbf{p}_k\} + \bar{u}_k - \bar{v}_k \\ & + c_k(\bar{\xi}_c - 1) \leq \bar{\xi}, \quad \forall k \in \mathcal{K} \end{aligned} \quad (19b)$$

$$\begin{aligned} & \mathbf{p}_c^H \bar{\Psi}_{c,k} \mathbf{p}_c + \sum_{i=1}^K \mathbf{p}_i^H \bar{\Psi}_{c,k} \mathbf{p}_i + \sigma_n^2 \bar{t}_{c,k} - 2\Re\{\bar{\mathbf{f}}_{c,k}^H \mathbf{p}_c\} \\ & + \bar{u}_{c,k} - \bar{v}_{c,k} \leq \bar{\xi}_c, \quad \forall k \in \mathcal{K} \end{aligned} \quad (19c)$$

$$\text{tr}(\mathbf{P} \mathbf{P}^H) \leq P_t. \quad (19d)$$

Problem (19) is convex with quadratic constraints, which can be solved using off-the-self optimization software that employs interior-point methods [10].

D. Alternating Optimization Algorithm

The AO algorithm is constructed by repeating the steps described in the previous subsections until convergence. This is summarized in Algorithm 2 where ϵ_R determines the

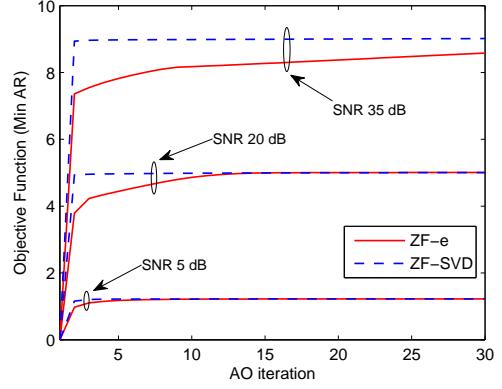


Fig. 1. Convergence of Algorithm 2 using 1 randomly generated \mathbf{H} .

accuracy of the solution and n_{\max} is the maximum number of iterations. Initializing \mathbf{P} is discussed in Section V.

Algorithm 2 Alternating Optimization

- 1: **Initialize:** $n \leftarrow 0, \hat{R}^{(n)} \leftarrow 0, \mathbf{P}$
- 2: **repeat**
- 3: $n \leftarrow n + 1, \ddot{\mathbf{P}} \leftarrow \mathbf{P}$
- 4: $\mathbf{G} \leftarrow \mathbf{G}^{\text{MMSE}}(\ddot{\mathbf{P}}), \mathbf{U} \leftarrow \mathbf{U}^{\text{MMSE}}(\ddot{\mathbf{P}})$
- 5: $\mathbf{c} \leftarrow \arg \mathcal{A}_{\mathbf{c}}$
- 6: update $\{\bar{\Psi}_{c,k}, \bar{\Psi}_k, \bar{\mathbf{F}}_{c,k}, \bar{\mathbf{F}}_k, \bar{t}_{c,k}, \bar{t}_k, \bar{u}_{c,k}, \bar{u}_k, \bar{v}_{c,k}, \bar{v}_k\}_{k=1}^K$
- 7: $(\mathbf{P}, \bar{\xi}^{(n)}) \leftarrow \arg \mathcal{A}_{\mathbf{P}}$ and $\hat{R}^{(n)} \leftarrow 1 - \bar{\xi}^{(n)}$
- 8: **until** $|\hat{R}^{(n)} - \hat{R}^{(n-1)}| < \epsilon_R$ or $n = n_{\max}$

Proposition 1. Algorithm 2 monotonically increases the AR objective function of problem \mathcal{R} until convergence.

Proof: From the steps in Algorithm 2, we observe the sequence: $1 - \bar{\xi}^{(n-1)} \leq \ddot{R} \leq \bar{R} \leq 1 - \bar{\xi}^{(n)}$, where $\ddot{R} = \min_k \{\ddot{c}_k \ddot{R}_c + \ddot{R}_k\}_{k=1}^K$ is the AR objective function at the output of step 4, and $\bar{R} = \min_k \{c_k \bar{R}_c + \bar{R}_k\}_{k=1}^K$ is the AR objective function after updating \mathbf{c} in step 5. The relationships in (11), (13) and (14) hold, particularly at the output of steps 4 and 5. This implies that each iteration increases the AR objective function. Moreover, the fact that \bar{R} is bounded above for given P_t ensures convergence. However, global optimality is not guaranteed due to the non-convexity of \mathcal{R} . ■

V. NUMERICAL RESULTS

We consider a system with $N_t = K = 2$ and uncorrelated channel fading, where the entries of \mathbf{H} have a complex Gaussian distribution $\mathcal{CN}(0, 1)$. The noise variance is fixed as $\sigma_n^2 = 1$, yielding a long-term SNR of P_t . Moreover, Gaussian CSIT error is assumed where the entries of $\tilde{\mathbf{H}}$ are generated according to the distribution $\mathcal{CN}(0, \sigma_e^2)$. For each realization \mathbf{H} , a channel estimation error $\tilde{\mathbf{H}}$ is drawn from $\mathcal{CN}(0, \sigma_e^2)$, from which the channel estimate is calculated as $\hat{\mathbf{H}} = \mathbf{H} - \tilde{\mathbf{H}}$. A channel realization \mathbf{H} should not be confused with a Monte-Carlo realization $\mathbf{H}^{(m)}$. The former is the actual channel which is unknown to the BS, and assumed to remain constant for a given transmission. On the other hand, the latter is part of a sample $\mathbf{H}_{\mathcal{M}}$ employed in the optimization at the BS. The size of the sample is set to $M = 1000$. For a given $\hat{\mathbf{H}}$, the

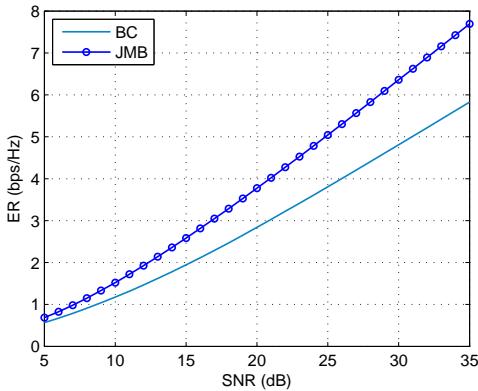


Fig. 2. MISO-BC and MISO-JMB ERs. $K = 2$, and $\sigma_e^2 = P_t^{-0.6}$.

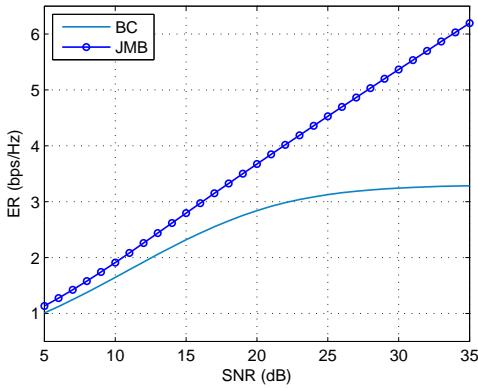


Fig. 3. MISO-BC and MISO-JMB ERs. $K = 2$, and $\sigma_e^2 = 0.063$.

m th Monte-Carlo realization is obtained as $\mathbf{H}^{(m)} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}^{(m)}$, where $\widetilde{\mathbf{H}}^{(m)}$ is drawn from $\mathcal{CN}(0, \sigma_e^2)$.

First, we examine the convergence of Algorithm 2 using two different \mathbf{P} initializations. The first initialization (ZF-e) is taken as the sum-DoF motivated design in [7]. The private precoders are initialized as $\mathbf{p}_k = \sqrt{P_t^\alpha / K} \widehat{\mathbf{p}}_k^{\text{ZF}}$, where $\widehat{\mathbf{p}}_k^{\text{ZF}}$ is a normalized ZF-BF vector constructed using the channel estimate $\widehat{\mathbf{H}}$. The common precoder is given as $\mathbf{p}_c = \sqrt{P_t - P_t^\alpha} \mathbf{e}_1$, where \mathbf{e}_1 is a vector with 1 as the first entry and zeros elsewhere. The second initialization (ZF-SVD) retains the ZF-BF part and the power allocation. However, the common precoder is obtained as the dominant left singular vector of $\widehat{\mathbf{H}}$. Figure 1 shows the AR convergence for $\sigma_e^2 = P_t^{-0.6}$, and SNRs 5, 20 and 35 dB. It is evident that the algorithm eventually converges to a limit point regardless of the initialization. Moreover, better convergence performance is achieved with ZF-SVD compared to ZF-e due to the initialization of the common precoder. The following results employ ZF-SVD.

Next, we consider the Ergodic Rate (ER) performance. For a given channel estimate $\widehat{\mathbf{H}}$ and the corresponding error statistics, maximizing the minimum AR yields the solution \mathbf{P} . Employing \mathbf{P} for the channel realization \mathbf{H} yields the achievable rates defined in II-C, from which the achievable minimum rate among users (i.e. R) is obtained. The ER is defined as $E_{\mathbf{H}}\{R\}$, which captures the overall achievable performance for all possible channel realizations. In the fol-

lowing simulations, the ER is calculated by averaging over 200 randomly generated channel realizations. Conventional fairness-based BC transmission is considered as a baseline, which is obtained by discarding the common ARs and \mathbf{c} in \mathcal{R} . A decaying estimation error power of $\sigma_e^2 = P_t^{-0.6}$ is used in Figure 2. On the other hand, Figure 3 uses a fixed estimation error power $\sigma_e^2 = (10^{\frac{20}{10}})^{-0.6} = 0.063$, which is equivalent to the CSIT quality obtained at SNR 20 dB in the previous case. It is evident that users benefit from the incorporation of the common message over the entire SNR range. The significance of this utility grows with increased SNR. For example, Figure 2 is a manifestation of DoF gains translating into rate gains at finitely high SNRs. Moreover, Figure 3 demonstrates JMB's virtue of achieving fairness with non-vanishing growth in a scenario where conventional transmission hits a performance ceiling.

VI. CONCLUSION

A multicast assisted scheme was proposed to achieve max-min fairness in a MU-MISO system with partial CSIT and perfect CSIR. Precoders are designed and the common message rate is divided among users, such that the minimum AR among users is maximized. This problem is transformed into an AWMSE problem solved using a converging AO algorithm. Simulation results show that the proposed scheme achieves significant rate gains over conventional transmission, which does not incorporate the multicast part.

REFERENCES

- [1] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan 2006.
- [2] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of Service and Max-Min Fair Transmit Beamforming to Multiple Cochannel Multicast Groups," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, March 2008.
- [3] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving maxmin fairness," *Signal Processing*, vol. 93, no. 12, pp. 3327 – 3340, 2013.
- [4] N. Vucic, H. Boche, and S. Shi, "Robust Transceiver Optimization in Downlink Multiuser MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3576–3587, Sept 2009.
- [5] T. Bogale, B. Chalise, and L. Vandendorpe, "Robust Transceiver Optimization for Downlink Multiuser MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 446–453, Jan 2011.
- [6] H. Joudeh and B. Clerckx, "AMMSE optimization for multiuser MISO systems with imperfect CSIT and perfect CSIR," in *IEEE Global Communications Conference (GLOBECOM)*, Austin, USA, Dec. 2014, pp. 3410–3415.
- [7] C. Hao and B. Clerckx, "MISO Broadcast Channel with imperfect and (Un)matched CSIT in the frequency domain: DoF region and transmission strategies," in *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 1–6.
- [8] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO Achievable Rates With Downlink Training and Channel State Feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.
- [9] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, December 2008.
- [10] M. Grant, S. Boyd, and Y. Ye, "CVX: MATLAB software for disciplined convex programming [Online]," Available: <http://www.stanford.edu/~boyd/cvx>, 2008.